UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/695,979	10/29/2003	Raimo Bakis	YOR9200379US1	8009

48175 7590 07/19/2007

BMT/IBM
50 LOCUST AVENUE
NEW CANAAN, CT 06840

EXAMINER	
GODBOLD, DOUGLAS	
ART UNIT	**PAPER NUMBER**
2626	
MAIL DATE	**DELIVERY MODE**
07/19/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/695,979	BAKIS ET AL.	
	Examiner	Art Unit	
	Douglas C. Godbold	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *25 May 2007*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-40* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-40* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This office action is in response to correspondence filed May 25, 2007 in reference to application 10/695,979. Claims 1- 40 are pending in the application and have been examined.

Response to Amendment

2. The amendments to the claims filed May 25, 2007 have been accepted and have been examined in this office action.

Response to Arguments

3. Applicant's arguments with respect to claims 1-40 have been considered but are moot in view of the new ground(s) of rejection.

Claim Rejections - 35 USC § 103

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

5. The factual inquiries set forth in *Graham* **v.** *John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under 35 U.S.C. 103(a) are summarized as follows:

1. Determining the scope and contents of the prior art.

2. Ascertaining the differences between the prior art and the claims at issue.
3. Resolving the level of ordinary skill in the pertinent art.
4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

6. Claims 1-33, 36, 37, and 40 are rejected under 35 U.S.C. 103(a) as being unpatentable over Henton (US Patent 5,860,064) in view of Kochanski et al (US Patent 6,810,378).

7. Consider claim 1, Henton teaches a method (figure 5), comprising:

identifying text to convert to speech (select text, step 501);

selecting a speech style sheet from a set of available speech style sheets, said speech style sheet defining desired speech characteristics (Choose vocal emotion for selected text; step 503);

marking said text to associate said text with said selected speech style sheet (figures 2-4 show marking text with colors, size, and boldface in order to associate text with a speech style); and

converting said text to speech having said desired speech characteristics by applying a low level markup generated by said speech style sheet (Look up synthesizer values for chosen emotion in emotion table [table 2], step 505. Apply speech synthesizer vocal emotion values to the chosen text, step 507.).

But Henton does not specifically teach wherein said selected speech style sheet defines pronunciation rules for a speech category and wherein another speech style sheet from said set of available speech style sheets defines pronunciation rules for another speech category.

In the same field of Speech Synthesizers, Kochanski teaches speech style sheet defines pronunciation rules for a speech category and wherein another speech style sheet from said set of available speech style sheets defines pronunciation rules for another speech category (It would be highly desirable to be able to capture a particular style, such as, for example, the style of a specifically identifiable person or of a particular class of people (e.g., a southern accent); column 1, line 28. In accordance with one illustrative embodiment of the present invention, a personal style for speech may be advantageously conveyed by repeated patterns of one or more features such as pitch, amplitude, spectral tilt, and/or duration, occurring at certain characteristic locations. These locations reflect the organization of speech materials. For example, a speaker may tend to use the same feature patterns at the end of each phrase, at the beginning, at emphasized words, or for terms newly introduced into a discussion column 2, line 53. Next, prosody evaluation module 55 converts the tags into a time series of prosodic features (or the equivalent) which can be used to directly control the synthesizer. The result of prosody evaluation module 55 may be referred to as a "stylized voice control information stream," since it provides voice control information adjusted for a particular style; column 5 line 15.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speaking styles that include accents which include pronunciation information of Kochanski with the style sheets of Henton in order to provide a more robust and flexible speech synthesis device.

8. Consider claim 2, Henton teaches a method according to claim 1, further comprising:

sending said text with said low level markup to an output device (Obtained vocal parameters will be outputted by the text to speech system; column 4, line 45. Values shown in Table 2 are input to the speech synthesizer, Column 10, line 42.).

9. Consider claim 3, Henton teaches a method according to claim 1, further comprising:

identifying at least one low level markup (columns of Table 2);

defining a voice style at least in part by associating said voice style with said at least one low level markup (Table 2 gives examples of the defined emotions of the preferred embodiment of the present invention with their associated vocal emotion values; column 9, line 56.); and

associating a speech style sheet with said voice style (Figure 1, device contains a memory for holding said vocal emotions parameters associated with emotions, column 4, line 54. Applicant defines the speech style sheet as a database; page 11, line 16. Therefore Henton teaches a style sheet.).

10. Consider claim 4, Henton teaches a method according to claim 3, wherein said associating said speech style sheet with said voice style includes:

creating said speech style sheet (As such, note that the particular values shown are easily modifiable, by the system implementer and/or the user, to thus allow for

differences in cultural interpretations and user/listener perceptions; column 9, line 61. If parameters are modifiable, one could easily create emotional styles.).

11. Consider claim 5, Henton teaches a method according to claim 3, wherein said associating said speech style sheet with said voice style includes:

editing said speech style sheet (As such, note that the particular values shown are easily modifiable, by the system implementer and/or the user, to thus allow for differences in cultural interpretations and user/listener perceptions; column 9, line 61.).

12. Consider claim 6, Henton teaches a method according to claim 1, wherein said low level markup defines at least one of a pitch, a prosody, a voice quality, a duration, a tremor, a timbre, a speed, an intonation, a timing, a volume, and a pronunciation rule (Table 2 gives examples of the defined emotions of the preferred embodiment of the present invention with their associated vocal emotion values; column 9, line 56. Table 2, shows pitch mean, range, volume, and speaking rate.).

13. Consider claim 7, Henton teaches a method according to claim 1, further comprising:

providing said speech style sheet to at least one of a text-to-speech developer and a text-to-speech device (As such, note that the particular values shown are easily modifiable, by the system implementer and/or the user, to thus allow for differences in cultural interpretations and user/listener perceptions; column 9, line 61. Style sheets

must be presented to a developer to be modified. Obtained vocal parameters will be outputted by the text to speech system; column 4, line 45. Values shown in Table 2 are input to the speech synthesizer, Column 10, line 42.).

14. Consider claim 8, Henton teaches a method according to claim 1, further comprising:

compiling a library of speech style sheets. (Figure 1, device contains a memory for holding said vocal emotions parameters associated with emotions, column 4, line 54. The vocal parameters associated with an emotion was inherently programmed into memory.)

15. Consider claim 9, Henton teaches a method according to claim 1, further comprising:

identifying at least one low level markup (column 11 lines 28-35 show text marked up with low level parameters.);

associating a speech style sheet with said at least one low level markup (Column 11 lines 28-35 show text marked up with low level parameters that were a result of applying different vocal emotions [from table 2] to different portions of text; column 11, line 1.).

16. Consider claim 10, Henton teaches a method according to claim 1, wherein said speech style sheet is selected from a menu of available speech style sheets (Figure 2 shows at the top a menu of emotions.).

17. Consider claim 11, Henton teaches a method according to claim 1, wherein said marking of said text includes annotating said text with an annotation such as underlining, bolding, italicizing, highlighting, color-coding, coding, adding a symbol, a mark, or a design (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; column 9, line 7.).

18. Consider claim 12, Henton teaches a method according to claim 1, wherein said converting said text to speech includes:

identifying said low level markup associated with said speech style sheet (Column 11 lines 28-35 show text marked up with low level parameters that were a result of applying different vocal emotions [from table 2] to different portions of text; column 11, line 1.); and

converting said marking of said text to said low level markup (Figures 2-4, text is marked using color codes to determine an emotion; described in detail column 7 line 60-column 9 line 11. Figure 5, Look up synthesizer values for chosen emotion in emotion table [table 2], step 505. Apply speech synthesizer vocal emotion values to the chosen text, step 507. Final marked up text with emotion values shown in column 11, line 28-35.).

19. Consider claim 13, Henton teaches a method according to claim 1, wherein said marking of said text further associates said text with a voice style associated with said speech style sheet (Figures 2-4, text is marked using color codes to determine an emotion; described in detail column 7 line 60-column 9 line 11. Emotions and parameters are shown in table 2.).

20. Consider claim 14, Henton teaches a method according to claim 13, wherein said voice style represents at least one of an age, an educational level, an emotion, a feeling, a physical trait, a personality trait, and a speech category (Henton teaches a method for automatic application of vocal emotion parameters, abstract.).

21. Consider claim 15, Henton teaches a method according to claim 1, wherein said low level markup allows a text-to-speech developer to convey a certain amount of information using less text. (Column 11 lines 28-35 show text marked up with low level parameters that were a result of applying different vocal emotions [from table 2] to different portions of text; column 11, line 1. These low level parameters convey information using text to the synthesizer.)

22. Consider claim 16, Henton teaches a method according to claim 1, wherein said selecting is performed by a text-to-speech developer not having expertise in voice arts (What is needed, therefore, is an intuitive graphical interface for specification and

modification of vocal emotion of synthetic speech; column 2, line 36. Further, the present invention provides for the automatic specification of prosodic controls which create vocal emotional affect in synthetic speech produced with a concatenative speech synthesizer, column 2, line 64.).

23. Consider claim 17, Henton teaches a speech style sheet (Figure 1, device contains a memory for holding said vocal emotions parameters associated with emotions, column 4, line 54. Applicant defines the speech style sheet as a database; page 11, line 16. Therefore Henton teaches a style sheet.), comprising:

at least one voice style associated with at least one voice-type, said at least one voice style relating a high level markup of said voice-type to a low level markup of said voice-type (Device contains a memory for holding said vocal emotions parameters associated with emotions, column 4, line 54. Associations are shown in table 2. Figures 2-4 show marking up text using color coding, bolding, and font size to associate emotions with text for emotions; column 9, line 7.), said at least one voice style including a voice of a particular gender, said other voice style further including a voice style representing a voice of another gender (Table 2 values are for a female voice, for a male voice the table values are to be altered, column 10, line 1.)

However Henton does not specifically teach that the male and female voices have different accents.

In the same of speech synthesis, Kochanski teaches using different accents to simulate different voices (It would be highly desirable to be able to capture a particular style,

such as, for example, the style of a specifically identifiable person or of a particular class of people (e.g., a southern accent); column 1, line 28. In accordance with one illustrative embodiment of the present invention, a personal style for speech may be advantageously conveyed by repeated patterns of one or more features such as pitch, amplitude, spectral tilt, and/or duration, occurring at certain characteristic locations. These locations reflect the organization of speech materials. For example, a speaker may tend to use the same feature patterns at the end of each phrase, at the beginning, at emphasized words, or for terms newly introduced into a discussion column 2, line 53. Next, prosody evaluation module 55 converts the tags into a time series of prosodic features (or the equivalent) which can be used to directly control the synthesizer. The result of prosody evaluation module 55 may be referred to as a "stylized voice control information stream," since it provides voice control information adjusted for a particular style; column 5 line 15.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speaking styles that include accents of Kochanski with the style sheets of Henton in order to provide a more robust and flexible speech synthesis device.

24. Consider claim 18, Henton teaches the speech style sheet according to claim 17, wherein said high level markup of said voice-type is a text markup (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; columns 7 line 61 - 9, line 11.).

25. Consider claim 19, Henton teaches the speech style sheet according to claim 17, wherein said high level markup includes at least one of an underlining, a bolding, an italicizing, a highlighting, a color-coding, an annotation, a coding, and an application of at least one of a symbol, a mark, and a design (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; columns 7 line 61 - 9, line 11.).

26. Consider claim 20, Henton teaches the speech style sheet according to claim 17, wherein said low level markup of said voice-type includes code causing generation of speech having particular speech properties (Column 11 lines 28-35 show text marked up with low level parameters that were a result of applying different vocal emotions [from table 2] to different portions of text; column 11, line 1. Values shown in Table 2 are input to the speech synthesizer, Column 10, line 42.).

27. Consider claim 21, Henton teaches the speech style sheet according to claim 17, wherein said low level markup defines at least one of a pitch, a prosody, a voice quality, a duration, a tremor, a timbre, speed, an intonation, a timing, a volume, and a pronunciation rule (Table 2 gives examples of the defined emotions of the preferred embodiment of the present invention with their associated vocal emotion values; column 9, line 56. Table 2, shows pitch mean, range, volume, and speaking rate.).

28. Consider claim 22, Henton teaches the speech style sheet according to claim 17, wherein said at least one voice style represents style characteristics such as an age, an educational level, an emotion, a feeling, a physical trait, a personality trait, and a speech category (Henton teaches a method for automatic application of vocal emotion parameters, abstract.).

29. Consider claim 23, Henton teaches the speech style sheet according to claim 17, wherein said speech style sheet is at least one of a programming object, a programming module, a computer program, or a computer file (Figure 1, device contains a memory for holding said vocal emotions parameters associated with emotions, column 4, line 54. The parameters must be saved in a computer file or program object to be stored by memory.).

30. Consider claim 24, Henton teaches an apparatus (figure 1), comprising:

a processor having access to at least one speech style sheet (CPU 11, connected to memory 17. Memory holds vocal emotion parameters associated with emotions; column 4, line 54.), said at least one speech style sheet containing a definition of a voice style associated with a voice-type, and said definition relating a high level markup of said voice-type to a low level markup of said voice-type (Device contains a memory for holding said vocal emotions parameters associated with emotions, column 4, line 54. Associations are shown in table 2. Figures 2-4 show marking up text using color coding, bolding, and font size to associate emotions with

text for emotions; column 9, line 7.), wherein said processor is operative to convert said high level markup to said low level markup (Look up synthesizer values for chosen emotion in emotion table [table 2], step 505. Apply speech synthesizer vocal emotion values to the chosen text, step 507.);

a user interface device for applying said at least one voice style to text associated with said voice-type, said user interface being in communication with said processor (Figure 1, a keyboard 13, or other textual input device such as a write-on tablet or touch screen, provides input to the CPU/memory unit 11, as does input controller 15 which by way of example can be a mouse, a 2-D trackball, a joystick, etc.; column 5, line 22.); and

an output device connected to said processor for converting said text with said low level markup to speech (figure 1, output 21. Values shown in Table 2 are input to the speech synthesizer, Column 10, line 42.).

But Henton does not specifically teach wherein said selected speech style sheet defines pronunciation rules for a speech category and wherein another speech style sheet from said set of available speech style sheets defines pronunciation rules for another speech category.

In the same field of Speech Synthesizers, Kochanski teaches speech style sheet defines pronunciation rules for a speech category and wherein another speech style sheet from said set of available speech style sheets defines pronunciation rules for another speech category (It would be highly desirable to be able to capture a particular style, such as, for example, the style of a specifically identifiable person or of a

particular class of people (e.g., a southern accent); column 1, line 28. In accordance with one illustrative embodiment of the present invention, a personal style for speech may be advantageously conveyed by repeated patterns of one or more features such as pitch, amplitude, spectral tilt, and/or duration, occurring at certain characteristic locations. These locations reflect the organization of speech materials. For example, a speaker may tend to use the same feature patterns at the end of each phrase, at the beginning, at emphasized words, or for terms newly introduced into a discussion column 2, line 53. Next, prosody evaluation module 55 converts the tags into a time series of prosodic features (or the equivalent) which can be used to directly control the synthesizer. The result of prosody evaluation module 55 may be referred to as a "stylized voice control information stream," since it provides voice control information adjusted for a particular style; column 5 line 15.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speaking styles that include accents which include pronunciation information of Kochanski with the style sheets of Henton in order to provide a more robust and flexible speech synthesis device.

31. Consider claim 25, Henton teaches the apparatus of claim 24, wherein said processor includes at least one of a text-to-speech engine (The preferred manner in which this invention would be implemented is in the context of creating vocal emotions that may be associated with text that is to be read by a text-to-speech synthesizer; column 9, line 15.) and a text normalizer (a simple linear normalization is then

performed in the preferred embodiment of the present invention in order to translate the graphical modifications to the resulting vocal emotion effect; column 9, line 38).

32. Consider claim 26, Henton teaches the apparatus according to claim 24, wherein said low level markup defines at least one of a pitch, a prosody, a voice quality, a duration, a tremor, a timbre, a speed, an intonation, a timing, a volume, and a pronunciation rule (Table 2 gives examples of the defined emotions of the preferred embodiment of the present invention with their associated vocal emotion values; column 9, line 56. Table 2, shows pitch mean, range, volume, and speaking rate.).

33. Consider claim 27, Henton teaches the apparatus according to claim 24, wherein said high level markup includes at least one of an underlining, a bolding, an italicizing, a highlighting, a color-coding, an annotation, a coding, and an application of at least one of a symbol, a mark, and a design (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; columns 7 line 61 - 9, line 11.).

34. Consider claim 28, Henton teaches the apparatus according to claim 24, wherein said voice style represents at least one of an age, an educational level, an emotion, a feeling, a physical trait, a personality trait, and a speech category (Henton teaches a method for automatic application of vocal emotion parameters, abstract.).

35. Consider claim 29, Henton teaches a system (Figure 1), comprising:

a designer device for creating speech style sheets (As such, note that the particular values shown are easily modifiable, by the system implementer and/or the user, to thus allow for differences in cultural interpretations and user/listener perceptions; column 9, line 61. If parameters are modifiable, one could easily create emotional styles.);

a speech style sheet at least partially created by said designer device , said speech style sheet defining a voice style (Figure 1, device contains a memory for holding said vocal emotions parameters associated with emotions, column 4, line 54. Applicant defines the speech style sheet as a database; page 11, line 16. Therefore Henton teaches a style sheet.);

said at least one voice style including a voice of a particular gender, said other voice style further including a voice style representing a voice of another gender (Table 2 values are for a female voice, for a male voice the table values are to be altered, column 10, line 1.)

a text-to-speech device for receiving text associated with a voice-type (The preferred manner in which this invention would be implemented is in the context of creating vocal emotions that may be associated with text that is to be read by a text-to-speech synthesizer; column 9, line 15.), said text having a high level markup associated with said voice style (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; columns 7 line 61 - 9, line 11.), said text-to-speech device having access to said speech style sheet (CPU 11, connected to memory 17.

Memory holds vocal emotion parameters associated with emotions; column 4, line 54.) and also having:

a memory for storing computer executable code (figure 1, memory 17); and

a processor for executing the program code stored in memory (CPU 11), wherein the program code includes;

code to determine, by accessing said speech style sheet, a low level markup associated with said high level markup (Figure 5, Look up synthesizer values for chosen emotion in emotion table [table 2], step 505.); and

code to convert said high level markup of said text to said low level markup (Apply speech synthesizer vocal emotion values to the chosen text, step 507.); and

an output device for producing expressive speech using said text with said low level markup, said output device in communication with said text-to-speech device (figure 1, output 21. Values shown in Table 2 are input to the speech synthesizer, Column 10, line 42.)

However Henton does not specifically teach that the male and female voices have different accents.

In the same of speech synthesis, Kochanski teaches using different accents to simulate different voices (It would be highly desirable to be able to capture a particular style, such as, for example, the style of a specifically identifiable person or of a particular

class of people (e.g., a southern accent); column 1, line 28. In accordance with one illustrative embodiment of the present invention, a personal style for speech may be advantageously conveyed by repeated patterns of one or more features such as pitch, amplitude, spectral tilt, and/or duration, occurring at certain characteristic locations. These locations reflect the organization of speech materials. For example, a speaker may tend to use the same feature patterns at the end of each phrase, at the beginning, at emphasized words, or for terms newly introduced into a discussion column 2, line 53. Next, prosody evaluation module 55 converts the tags into a time series of prosodic features (or the equivalent) which can be used to directly control the synthesizer. The result of prosody evaluation module 55 may be referred to as a "stylized voice control information stream," since it provides voice control information adjusted for a particular style; column 5 line 15.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speaking styles that include accents of Kochanski with the style sheets of Henton in order to provide a more robust and flexible speech synthesis device.

36. Consider claim 30, Henton teaches the system according to claim 29, further comprising:

a developer device in communication with said text-to-speech device (Figure 1, a keyboard 13, or other textual input device such as a write-on tablet or touch screen, provides input to the CPU/memory unit 11, as does input controller 15 which by way of

example can be a mouse, a 2-D trackball, a joystick, etc.; column 5, line 22.), said developer device for marking text and providing said text to said text-to-speech device (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; columns 7 line 61 - 9, line 11.).

37. Consider claim 31, Henton teaches the system according to claim 29, further comprising:

a user interface device in communication with said text-to-speech device (Figure 1, a keyboard 13, or other textual input device such as a write-on tablet or touch screen, provides input to the CPU/memory unit 11, as does input controller 15 which by way of example can be a mouse, a 2-D trackball, a joystick, etc.; column 5, line 22.), said user interface device for applying high level markup to text and providing said text to said text-to-speech device (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; columns 7 line 61 - 9, line 11.).

38. Consider claim 32, Henton teaches an article of manufacture (figure 1), comprising:

a computer usable medium having computer readable program code means embodied therein for producing expressive text-to-speech (External storage 17, which can include fixed disk drives, floppy disk drives, memory cards, etc., is used for mass storage of programs and data; column 5, line 26. Method, figure 5.), comprising:

computer readable program code means for identifying text to convert to speech (select text, step 501);

computer readable program code means for selecting a speech style sheet from a set of available speech style sheets, said speech style sheet defining desired speech characteristics (Choose vocal emotion for selected text; step 503);

computer readable program code means for marking said text to associate said text with said selected speech style sheet (figures 2-4 show marking text with colors, size, and boldface in order to associate text with a speech style); and

computer readable program code means for converting said text to speech having said desired speech characteristics by applying a low level markup associated with said speech style sheet (Look up synthesizer values for chosen emotion in emotion table [table 2], step 505. Apply speech synthesizer vocal emotion values to the chosen text, step 507.).

But Henton does not specifically teach wherein said selected speech style sheet defines pronunciation rules for a speech category and wherein another speech style sheet from said set of available speech style sheets defines pronunciation rules for another speech category.

In the same field of Speech Synthesizers, Kochanski teaches speech style sheet defines pronunciation rules for a speech category and wherein another speech style sheet from said set of available speech style sheets defines pronunciation rules for another speech category (It would be highly desirable to be able to capture a particular

style, such as, for example, the style of a specifically identifiable person or of a particular class of people (e.g., a southern accent); column 1, line 28. In accordance with one illustrative embodiment of the present invention, a personal style for speech may be advantageously conveyed by repeated patterns of one or more features such as pitch, amplitude, spectral tilt, and/or duration, occurring at certain characteristic locations. These locations reflect the organization of speech materials. For example, a speaker may tend to use the same feature patterns at the end of each phrase, at the beginning, at emphasized words, or for terms newly introduced into a discussion column 2, line 53. Next, prosody evaluation module 55 converts the tags into a time series of prosodic features (or the equivalent) which can be used to directly control the synthesizer. The result of prosody evaluation module 55 may be referred to as a "stylized voice control information stream," since it provides voice control information adjusted for a particular style; column 5 line 15.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speaking styles that include accents which include pronunciation information of Kochanski with the style sheets of Henton in order to provide a more robust and flexible speech synthesis device.

39. Consider claim 33, Henton teaches a system for producing expressive text-to-speech, (system figure 1, Method figure 5), comprising:

means for identifying text to convert to speech (select text, step 501);

means for selecting a speech style sheet from a set of available speech style sheets, said speech style sheet defining desired speech characteristics (Choose vocal emotion for selected text; step 503);

means for marking said text to associate said text with said selected speech style sheet (figures 2-4 show marking text with colors, size, and boldface in order to associate text with a speech style); and

means for converting said text to speech having said desired speech characteristics by applying a low level markup associated with said speech style sheet (Look up synthesizer values for chosen emotion in emotion table [table 2], step 505. Apply speech synthesizer vocal emotion values to the chosen text, step 507.).

But Henton does not specifically teach wherein said selected speech style sheet defines pronunciation rules for a speech category and wherein another speech style sheet from said set of available speech style sheets defines pronunciation rules for another speech category.

In the same field of Speech Synthesizers, Kochanski teaches speech style sheet defines pronunciation rules for a speech category and wherein another speech style sheet from said set of available speech style sheets defines pronunciation rules for another speech category (It would be highly desirable to be able to capture a particular style, such as, for example, the style of a specifically identifiable person or of a particular class of people (e.g., a southern accent); column 1, line 28. In accordance with one illustrative embodiment of the present invention, a personal style for speech may be advantageously conveyed by repeated patterns of one or more features such as

pitch, amplitude, spectral tilt, and/or duration, occurring at certain characteristic locations. These locations reflect the organization of speech materials. For example, a speaker may tend to use the same feature patterns at the end of each phrase, at the beginning, at emphasized words, or for terms newly introduced into a discussion column 2, line 53. Next, prosody evaluation module 55 converts the tags into a time series of prosodic features (or the equivalent) which can be used to directly control the synthesizer. The result of prosody evaluation module 55 may be referred to as a "stylized voice control information stream," since it provides voice control information adjusted for a particular style; column 5 line 15.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speaking styles that include accents which include pronunciation information of Kochanski with the style sheets of Henton in order to provide a more robust and flexible speech synthesis device.

40. Claims 34, and 38 are rejected under 35 U.S.C. 103(a) as being unpatentable over Henton in view of Kochanski as applied to claims 1 and 24 above, and further in view of Atkin et al (US PAP 2004/0260551).

41. Consider claim 34, Henton in view of Kochanski teaches the method according to claim 1, but does not specifically teach wherein said selected speech style sheet defines pronunciation rules for at least one of aviation, chemistry and real estate. However in the same field of speech to text, Atkin suggests said selected speech style

sheet defines pronunciation rules for at least one of aviation, chemistry and real estate (A subject matter semantic identifier corresponds to particular subject matter, such as a children's book or a financial article. A user interest semantic identifier corresponds to particular areas of interest, such as a summary, detail, or section headings of a text file. For example, the semantic analyzer identifies that a text block is a paragraph corresponding to financial information and associates a "Business Journal" semantic identifier with the text block. In this example, the semantic analyzer retrieves voice attributes corresponding to the "Business Journal" semantic identifier from the look-up table. The semantic analyzer provides the voice attributes to a voice reader. The voice attributes include attributes such as a pitch value, a loudness value, and a pace value. In one embodiment, the voice attributes are provided to the voice reader through an Application Program Interface (API). The voice reader inputs the voice attributes into a voice synthesizer whereby the voice synthesizer converts the text block into synthesized speech for a user to hear; paragraphs 0010 and 0011. Although it does not specifically say aviation or chemistry or real estate, one of ordinary skill in the art could appreciate that this process is applicable to these fields as well.).

Therefore it would have been obvious to one of ordinary skill in the art to use the context dependency as taught by Atkin with the style sheets of Henton in view of Kochanski in order to provide a context dependent speech synthesizer.

42. Consider claim 38, Henton in view of Kochanski teaches the apparatus according to claim 24, but does not specifically teach wherein said selected speech style sheet

defines pronunciation rules for at least one of aviation, chemistry and real estate. However in the same field of speech to text, Atkin suggests said selected speech style sheet defines pronunciation rules for at least one of aviation, chemistry and real estate (A subject matter semantic identifier corresponds to particular subject matter, such as a children's book or a financial article. A user interest semantic identifier corresponds to particular areas of interest, such as a summary, detail, or section headings of a text file. For example, the semantic analyzer identifies that a text block is a paragraph corresponding to financial information and associates a "Business Journal" semantic identifier with the text block. In this example, the semantic analyzer retrieves voice attributes corresponding to the "Business Journal" semantic identifier from the look-up table. The semantic analyzer provides the voice attributes to a voice reader. The voice attributes include attributes such as a pitch value, a loudness value, and a pace value. In one embodiment, the voice attributes are provided to the voice reader through an Application Program Interface (API). The voice reader inputs the voice attributes into a voice synthesizer whereby the voice synthesizer converts the text block into synthesized speech for a user to hear; paragraphs 0010 and 0011. Although it does not specifically say aviation, one of ordinary skill in the art could appreciate that this process is applicable to these fields as well.).

Therefore it would have been obvious to one of ordinary skill in the art to use the context dependency as taught by Atkin with the style sheets of Henton in view of Kochanski in order to provide a context dependent speech synthesizer.

43. Consider claim 36, Henton teaches The speech style sheet according to claim 17, wherein said language is English (All examples in figures 204 are in English.)

44. Consider claim 37, Henton and Kochanski suggest the speech style sheet according to claim 17, wherein said particular gender is male (Henton, Table 2 values are for a female voice, for a male voice the table values are to be altered, column 10, line 1.), said language is common English (Henton, all examples in figures 2-4 are in English), said accent is a southern U.S. accent and said another accent is a Cornish accent (It would be highly desirable to be able to capture a particular style, such as, for example, the style of a specifically identifiable person or of a particular class of people (e.g., a southern accent); column 1, line 28. In accordance with one illustrative embodiment of the present invention, a personal style for speech may be advantageously conveyed by repeated patterns of one or more features such as pitch, amplitude, spectral tilt, and/or duration, occurring at certain characteristic locations. These locations reflect the organization of speech materials. For example, a speaker may tend to use the same feature patterns at the end of each phrase, at the beginning, at emphasized words, or for terms newly introduced into a discussion column 2, line 53. Next, prosody evaluation module 55 converts the tags into a time series of prosodic features (or the equivalent) which can be used to directly control the synthesizer. The result of prosody evaluation module 55 may be referred to as a "stylized voice control information stream," since it provides voice control information adjusted for a particular style; column 5 line 15. Although a Cornish accent is not specifically taught, it would be

obvious to one of ordinary skill in the art that one could be included in the available styles.)

45. Consider claim 40, Henton teaches The speech style sheet according to claim 29, wherein said language is English (All examples in figures 204 are in English.)

46. Claims 35 and 39 are rejected under 35 U.S.C. 103(a) as being unpatentable over Henton in view of Kochanski as applied to claims 1 and 24 above, and further in view of Surace et al (US Patent 6,334,103.).

47. Consider claim 35, Henton in view of Kochanski teaches the method according to claim 1, but does not teach specifically wherein said selected speech style sheet defines pronunciation rules for an automated flight reservation system.

In the same field of speech synthesis, Surace suggests said selected speech style sheet defines pronunciation rules for an automated flight reservation system. (In one embodiment, controlling the voice user interface includes providing the voice user interface with multiple personalities. The voice user interface with personality installs a prompt suite for a particular personality from a prompt repository that stores multiple prompt suites, in which the multiple prompt suites are for different personalities of the voice user interface with personality; column 2, line 12. Although this art does not specifically teach a flight reservation, one of ordinary skill in the art can appreciate that a prompting voice system can be used as a flight reservation system.)

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use a voice interface with personality as taught by Surace as an application for the style sheet system of Henton in view of Kochanski in order to provide a personalized experience in a voice response system.

48. Consider claim 39, Henton in view of Kochanski teaches the apparatus according to claim 24, but does not teach specifically wherein said selected speech style sheet defines pronunciation rules for an automated flight reservation system.

In the same field of speech synthesis, Surace suggests said selected speech style sheet defines pronunciation rules for an automated flight reservation system. (In one embodiment, controlling the voice user interface includes providing the voice user interface with multiple personalities. The voice user interface with personality installs a prompt suite for a particular personality from a prompt repository that stores multiple prompt suites, in which the multiple prompt suites are for different personalities of the voice user interface with personality; column 2, line 12. Although this art does not specifically teach a flight reservation, one of ordinary skill in the art can appreciate that a prompting voice system can be used as a flight reservation system.)

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use a voice interface with personality as taught by Surace as an application for the style sheet system of Henton in view of Kochanski in order to provide a personalized experience in a voice response system.

Conclusion

49. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Douglas C. Godbold whose telephone number is (571) 270-1451. The examiner can normally be reached on Monday-Thursday 7:00am-4:30pm Friday 7:00am-3:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick Edouard can be reached on (571) 272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DCG

PATRICK N. EDOUARD
SUPERVISORY PATENT EXAMINER